

长相依面板数据的斜率变点分析

朱 旭, 庞天晓

(浙江大学 数学科学学院, 浙江杭州 310058)

摘 要: 金融面板数据往往同时存在结构变点, 时间序列长相依以及横截面相依的现象. 假设时间序列长度为 T , 个体数为 N , 斜率变化发生在所有个体的同一时刻, 每个时间序列是长相依的(记忆参数 $d \in (0, 0.5)$), 个体之间通过公因子结构而具有横截面的相依性. 对于这种面板数据模型, 用最小二乘的方法估计斜率变点发生的时刻以及发生时刻的分数, 并研究了当 (T, N) 联合趋于无穷时估计量的渐近性质: 估计量的相合性, 收敛速度以及极限分布. 得到了一些有趣的结论: 大部分情况下斜率变点估计量的收敛速度随着记忆参数 d 的增大而减缓, 但当面板数据的时间长度 T 和个体数 N 满足 $T^{2d} = o(N)$, 且公因子与变点的变化幅度存在交互效应时, 变点时刻估计量的收敛速度仅与 T 有关. Monte Carlo模拟评估了估计量在有限样本情形下的表现, 并支持文中的理论结果.

关键词: 面板数据; 长相依; 斜率变点; 最小二乘估计

中图分类号: O212.1

文献标识码: A **文章编号:** 1000-4424(2022)04-0379-18

§1 引 言

许多经济和金融时间序列的数据都具有单变点或者多变点的特征, 参见Bai和Perron^[1], Hansen^[2], Lee等^[3], 以及Perron和Zhu^[4]. 此外众所周知, 经济和金融数据往往还具有长相依的特征. 例如日元兑美元的汇率数据被广泛认为存在长相依, 参见Horváth和Kokoszka^[5]. 因此在处理实际数据时, 往往面临结构变点和长相依同时存在的情形. 例如Jarušková^[6]在非平稳的长周期的水文数据中发现了长相依的证据, 同时认为数据中存在结构变点. 大量的研究表明, 长相依的存在会使传统的变点理论失效. 因此对长相依数据进行变点分析既具有理论意义, 也具有实际意义.

有相当多的文章研究了长相依时间序列中的结构变点. 例如Kuan和Hsu^[7]研究了均值变点并给出了估计量的收敛速度. Chang和Perron^[8]研究了参数为 d 的分整(fractionally integrated)过程中的斜率和均值变点, 并给出了估计量的收敛速度和极限分布. Perron和Qu^[9]提出用周期图

来研究股票指数回报的波动率,发现存在长相依和均值漂移. Shao^[10]提出一种检验流程来检验长相依时间序列中的均值变点. Iacone等作者^[11]提出一种检验流程来检验参数为 d 的分整时间序列中的斜率变点.

面板数据中存在公共变点的现象在实际生活中非常普遍. 例如税收政策的变化可能会影响每个人的收入. 另一个著名的例子是2008年发生的全球金融危机几乎影响了所有国家的GDP(Gross Domestic Product)数据和资本市场表现,并带来一定时间内持续的影响. 类似地,新的科学技术的诞生,新药的研发生产,以及政府层面新的政策措施都会对经济社会造成或多或少的影响.

结构变点从时间序列的研究延申到面板数据的研究可以追溯到1990年. Joseph和Wolfson是最早的一批研究面板数据中的结构变点的学者^[12-13],他们提出了 N 个面板上存在独立同分布的 N 个变点的随机变点模型,并给出了变点分布的相合估计量. 不久后随机变点模型被Joseph等^[14]推广到了自回归模型. Bai^[15]利用最小二乘的方法估计了面板数据中的公共均值变点,同时证实了变点估计量是相合的,这与平稳时间序列中变点估计量是不相合的结论截然相反. 值得一提的是,以上文章中的面板数据都是假设横截面独立的. 此后,面板数据中的公共变点研究得到了计量经济学界和统计学界的大量关注,横截面独立的面板数据模型也在一些文章中被推广到了横截面相依情形. 例如,Horváth和Hušková^[16]研究了存在横截面相依的面板数据中公共均值变点的检验问题, Kim^[17-18]研究了横截面相依的面板数据中存在斜率变点以及斜率与截距变点同时存在时的变点估计问题, Li等作者^[19]利用自适应group fused LASSO惩罚的主成分方法估计横截面相依的面板数据中的变点, Qian和Su^[20]提出了横截面相依面板数据中的公共变点的收缩估计量, Baltagi等作者^[21-22]提出了CCE(common correlated effects)估计量并把它应用于横截面相依面板数据中变点的研究, Westerlund^[23]采用了基于CCE的方法对时间长度固定的横截面相依面板数据中的变点进行研究. 值得注意的是,大部分现有的文献都假设面板数据在时间维度上是弱相依的,研究长相依面板数据中的结构变点的文章相对较少.

本文假设面板数据具有横截面的相依性,同时每个时间序列都是长相依的(记忆参数 $d \in (0, 0.5)$)且带有斜率变点. 这种模型在宏观经济和金融中具有广泛的应用. 用最小二乘的方法估计斜率变点发生的时刻以及发生时刻的分数,并研究了当 (T, N) (T 表示时间长度, N 表示个体数)联合趋于无穷时估计量的渐近性质: 估计量的相合性,收敛速度以及极限分布. 得到了一些有趣的结论: 大部分情况下斜率变点估计量的收敛速度随着记忆参数 d 的增大而减缓,但当 T 和 N 满足 $T^{2d} = o(N)$,且公因子与变点的变化幅度存在交互效应时,变点时刻估计量的收敛速度仅与 T 有关.

本文结构如下: §2给出模型的具体形式以及一些假设, §3介绍变点估计量的渐近性质, §4给出Monte Carlo模拟结果,用于评估变点估计量在有限样本情形下的表现, §5对论文进行了总结,文章中的理论证明放在§6.

§2 模型与假设

首先介绍本文接下来需要用到的一些记号. “ \Rightarrow ”表示概率测度弱收敛(参见Billingsley^[24]), “ \xrightarrow{P} ”表示依概率收敛, “ \xrightarrow{d} ”表示依分布收敛, “iid”表示独立同分布, $a_T \asymp b_T$ 表示存在两个正常数 c_1 和 c_2 使得对于所有足够大的 T , $c_1 \leq a_T/b_T \leq c_2$ 成立,其中 a_T 和 b_T 是两个关于 T 的取值恒为正的函数. 设 $\mathcal{D}[0, 1]$ 是 $[0, 1]$ 上的函数空间,且是存在左极限和右连续的Skorohod拓扑. 对于矩

阵 A , 记 $\|A\| = [\text{tr}(A'A)]^{1/2}$, 其中 $\text{tr}[\cdot]$ 表示矩阵的迹. 此外用 M 表示正常数, 它的取值与 T 和 N 无关, 但在不同的地方可取不同的值. 除非另有说明, 本文中的极限均理解为 $(T, N) \rightarrow \infty$, 即 T 和 N 联合趋于无穷.

本文研究的面板数据模型为

$$y_{i,t} = \mu_i + \beta_i t + \gamma_i B_{T_1}(t) + u_{i,t}, \quad t = 1, \dots, T; \quad i = 1, \dots, N,$$

其中 $B_{T_1}(t)$ 是关于斜率变化的哑变量, 其定义为

$$B_{T_1}(t) = \begin{cases} 0, & t \leq T_1, \\ t - T_1, & t > T_1, \end{cases}$$

μ_i 表示截距, 斜率参数在时刻 T_1 从 β_i 变化到 $\beta_i + \gamma_i$. Kim^[17]曾在 $I(0)$ 和 $I(1)$ 的模型误差假设下研究过该模型.

为了简化表述, 接下来考虑模型的矩阵表示. 记

$$Y_i = \mathbf{X}'_{T_1} \Pi_i + U_i,$$

以及

$$\mathbf{Y} = \Pi \mathbf{X}_{T_1} + \mathbf{U},$$

这里 $Y_i = (y_{i,1}, \dots, y_{i,T})'$, $\mathbf{Y} = (Y_1, \dots, Y_N)'$, $U_i = (u_{i,1}, \dots, u_{i,T})'$, $\mathbf{U} = (U_1, \dots, U_N)'$, $\mathbf{X}_{T_1} = (\iota, \tau, B_{T_1})' = (x(T_1)_1, \dots, x(T_1)_T)$, 其中 $\iota = (1, \dots, 1)'$, $\tau = (1, 2, \dots, T)'$, $B_{T_1} = (B_{T_1}(1), \dots, B_{T_1}(T))'$, $x(T_1)_t = (1, t, B_{T_1}(t))'$, 以及 $\Pi = (\Pi_1, \dots, \Pi_N)' = (\mu, \beta, \gamma)$, 其中 $\mu = (\mu_1, \dots, \mu_N)'$, $\beta = (\beta_1, \dots, \beta_N)'$, $\gamma = (\gamma_1, \dots, \gamma_N)'$. 注意这里的矩阵 \mathbf{X}_{T_1} 中的 T_1 是一个假定的变点时刻. 此外假设

$$u_{i,t} = h'_i F_t + e_{i,t},$$

其中 F_t 是一个 $m \times 1$ 的公因子向量, h_i 是 $m \times 1$ 的因子载荷, $e_{i,t}$ 是误差. 进一步写成矩阵形式有

$$Y_i = \mathbf{X}'_{T_1} \Pi_i + \mathbf{F}' h_i + E_i,$$

其中

$$\mathbf{F} = (F_1, \dots, F_T), \quad E_i = (e_{i,1}, \dots, e_{i,T})'.$$

再记

$$\mathbf{Y} = \Pi \mathbf{X}_{T_1} + \mathbf{H} \mathbf{F} + \mathbf{E},$$

其中 $\mathbf{H} = (h_1, \dots, h_N)'$, $\mathbf{E} = (E_1, \dots, E_N)'$.

对于如上模型, 利用最小二乘的方法去估计变点时刻. T_1 的最小二乘估计量定义为

$$\hat{T}_1 = \arg \min_{T_1 \in [\pi T, (1-\pi)T]} \text{tr}[\mathbf{Y}(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{Y}'], \quad (1)$$

其中 π 是一个很小的正常数(在实际应用中可以取 $\pi = 0.05$), \mathbf{I} 是单位矩阵, \mathbf{P}_{T_1} 是由 \mathbf{X}_{T_1} 构成的投影矩阵, 即 $\mathbf{P}_{T_1} = \mathbf{X}'_{T_1} (\mathbf{X}_{T_1} \mathbf{X}'_{T_1})^{-1} \mathbf{X}_{T_1}$. 因此变点分数的估计量被定义为

$$\hat{\lambda} = \hat{T}_1 / T.$$

以下约定: 所有包含上标0的参数都代表其真值. 变点时刻表示为 T_1 , 其真值为 T_1^0 . 同理变点分数 $\lambda = T_1 / T$, 其真值为 λ^0 .

此外当 $T_1 > T_1^0$ 时, 定义

$$\tilde{b}(t; T_1) = \begin{cases} 0, & 1 \leq t \leq T_1^0, \\ \frac{t - T_1^0}{T_1 - T_1^0}, & T_1^0 < t \leq T_1, \\ 1, & T_1 < t \leq T. \end{cases}$$

当 $T_1 = T_1^0$ 时定义

$$\tilde{t}_b(t; T_1^0) = \begin{cases} 0, & 1 \leq t \leq T_1^0, \\ 1, & T_1^0 < t \leq T. \end{cases}$$

当 $T_1 < T_1^0$ 时定义

$$\tilde{t}_b(t; T_1) = \begin{cases} 0, & 1 \leq t \leq T_1, \\ \frac{t-T_1}{T_1^0-T_1}, & T_1 < t \leq T_1^0, \\ 1, & T_1^0 < t \leq T. \end{cases}$$

记 $\tilde{t}_b = (\tilde{t}_b(1; T_1), \dots, \tilde{t}_b(T; T_1))$.

接下来介绍面板数据在时间维度上的长相依(参见Wang等作者^[25]). 定义

$$(1-L)^a = \sum_{j=0}^{\infty} \pi_j(a) L^j, \quad \pi_j(a) = \frac{\Gamma(j-a)}{\Gamma(-a)\Gamma(j+1)},$$

其中 L 是时间维度上的滞后算子, $\Gamma(\cdot)$ 是Gamma函数. 若对 $d \in (0, 0.5)$,

$$(1-L)^d \xi_t = \omega_t, \quad t = 0, \pm 1, \dots,$$

这里 $\{\omega_t, t = 0, \pm 1, \dots\}$ 是一个iid序列或者弱相依的平稳序列, 则称 $\{\xi_t, t = 0, \pm 1, \dots\}$ 是记忆参数为 d 的长相依随机变量序列.

为了后面的极限性质的推导, 对模型提出如下假设.

假设1 真实变点时刻 T_1^0 是未知的, 且存在一个较小的 π , $0 < \pi < 1/2$, 变点分数满足 $\lambda^0 \in (\pi, 1-\pi) \subset (0, 1)$.

假设2

(1) 公因子向量 F_t 满足 $F_t = \psi(L)\epsilon_t$, 这里 $\epsilon_t \sim \text{iid}(0, \mathbf{I}_m)$, $\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$, 其中 ψ_j 是一个 $m \times m$ 的矩阵. 此外, $\sum_{j=0}^{\infty} j \|\psi_j\| < M$, $\det(\psi(z)) \neq 0$.

(2) 对于 $d \in (0, 0.5)$, $(1-L)^d e_{i,t} = \varepsilon_{i,t}$, 其中 $\varepsilon_{i,t} = \sigma_i \eta_{i,t}$, $\eta_{i,t}$ 在 i 和 t 上都满足 $\eta_{i,t} \sim \text{iid}(0, 1)$ 且 $\sigma_i^2 \leq M$.

(3) $[\tilde{t}_b(\mathbf{I} - \mathbf{P}_{T_1})E_i/\sigma_i]^2$ 关于 T 一致可积.

(4) \mathbf{F} 和 \mathbf{E} 相互独立.

假设3 $N^{-1}\gamma'\gamma \rightarrow A_{\gamma\gamma} \neq 0$ 且 $N^{-1}\gamma'\Sigma_\varepsilon\gamma \rightarrow S_{\gamma\gamma} \neq 0$, 这里 $\Sigma_\varepsilon = \text{diag}\{\sigma_1^2, \dots, \sigma_N^2\}$. 此外, $\max\{\gamma_1^2, \dots, \gamma_N^2\} = O(1)$.

假设4 h_i 和 γ_i 满足 $N^{-1}\mathbf{H}'\gamma \rightarrow A_{H\gamma}$ 以及 $N^{-1}\mathbf{H}'\mathbf{H} \rightarrow \mathbf{A}_{HH}$, 这里 $A_{H\gamma}$ 是一个固定的向量, \mathbf{A}_{HH} 是一个固定的矩阵.

注 假设1是为了保证变点的可识别性, 是变点研究领域中的常见假设. 假设2(1)假设公因子是一个线性过程并可以应用泛函中心极限定理(参见Phillips和Solo^[26]). 假设2(2)假设误差 $e_{i,t}$ 在时间维度上是长相依的, 但其方差在横截面维度上可以不同. 假设2(3)是一个技术性假设, 是为了应用面板数据的联合中心极限定理(参见Phillips和Moon^[27]). 假设2(4)假设公因子和模型误差是相互独立的. 假设3和4描述了模型的总体变点强度, 公因子的强度, 以及变点的变化幅度与公因子之间的相互影响程度.

§3 极限性质

本节介绍估计量 \hat{T}_1 的极限性质. 后文中, 所有的 $O_p(\cdot)$ 将按照其严格的定义理解, 即随机变

量不会是 $o_p(\cdot)$.

定理3.1 若假设1-4成立, 当 $(T, N) \rightarrow \infty$, $d \in (0, 0.5)$ 时, 有下面的结论.

- (1) 若 $N = o(T^{3-2d})$ 或 $N \asymp T^{3-2d}$, 且对所有 i 都有 $h_i = 0$, 则 $|\hat{T}_1 - T_1^0| = O_p(T^{-1/2+d}N^{-1/2})$.
- (2) 若 $N = o(T^{2d})$ 或 $N \asymp T^{2d}$, 且 $A_{H\gamma} \neq 0$, 则 $|\hat{T}_1 - T_1^0| = O_p(T^{-1/2+d}N^{-1/2})$.
- (3) 若 $T^{2d} = o(N)$, 且 $A_{H\gamma} \neq 0$, 则 $|\hat{T}_1 - T_1^0| = O_p(T^{-1/2})$.

注 以上结论都表明变点估计量 \hat{T}_1 是相合的. 结论(1)说明在不存在横截面相依时, 收敛速度与记忆参数 d , T 和 N 有关, 且 d 越大收敛速度越慢, 这与直觉吻合, 因为 $d \geq 0.5$ 时时间序列将变成非平稳. 结论(2)说明当 $A_{H\gamma} \neq 0$ 时, 即公因子与变点的变化幅度存在交互效应, 如果个体数 N 相对于时间长度 T 并没有足够大时, 收敛速度依然由 d , T 和 N 共同决定, 且随着 d 的增大而变慢. 结论(3)说明当 $T^{2d} = o(N)$, 即个体数 N 相对于时间长度 T 足够大时, \hat{T}_1 的收敛速度仅与 T 有关.

有了收敛速度, 接下来推导估计量的极限分布. 有下面的定理.

定理3.2 若假设1-4成立, 当 $(T, N) \rightarrow \infty$, $d \in (0, 0.5)$ 时, 有下面的结论.

- (1) 若对所有 i 都有 $h_i = 0$, 且 $N = o(T^{3-2d})$; 或者 $A_{H\gamma} \neq 0$, 且 $N = o(T^{2d})$, 则

$$T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0) \xrightarrow{d} N\left(0, \frac{16\tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2 S_{\gamma\gamma}}{(1-\lambda^0)^2(\lambda^0)^2 A_{\gamma\gamma}^2}\right),$$

这里 $\tilde{\kappa}(d)^2 = \{\Gamma(1-2d)\}/\{(1+2d)\Gamma(1+d)\Gamma(1-d)\}$, $\sigma_{\tilde{\eta}}^2 = \text{Var}(\zeta_{\tilde{\eta}})$,

$$\zeta_{\tilde{\eta}} = \int_0^{\lambda^0} \frac{\lambda^0 - (\lambda^0)^2 - 3r + 3r\lambda^0}{2\lambda^0} dB_d(r) + \int_{\lambda^0}^1 \frac{\lambda^0(2 + \lambda^0 - 3r)}{2(1-\lambda^0)} dB_d(r),$$

$B_d(\cdot)$ 是 $\mathcal{D}[0, 1]$ 上的分数布朗运动, 即

$$B_d(t) = \frac{1}{\Gamma(d+1)} \int_{-\infty}^0 [(t-s)^d - (-s)^d] dB(s) + \int_0^t (t-s)^d dB(s),$$

其中 $B(\cdot)$ 为标准布朗运动.

- (2) 若 $A_{H\gamma} \neq 0$, 且 $N/T^{2d} \rightarrow c \neq 0$, 则

$$T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0) \xrightarrow{d} N\left(0, \frac{4cA'_{H\gamma}\psi(1)\psi(1)'A_{H\gamma}}{\lambda^0(1-\lambda^0)A_{\gamma\gamma}^2} + \frac{16\tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2 S_{\gamma\gamma}}{(1-\lambda^0)^2(\lambda^0)^2 A_{\gamma\gamma}^2}\right).$$

- (3) 若 $A_{H\gamma} \neq 0$, 且 $T^{2d} = o(N)$, 则

$$T^{1/2}(\hat{T}_1 - T_1^0) \xrightarrow{d} N\left(0, \frac{4A'_{H\gamma}\psi(1)\psi(1)'A_{H\gamma}}{\lambda^0(1-\lambda^0)A_{\gamma\gamma}^2}\right).$$

注 以上的三个结论表明: 当变点分数的真值 λ^0 越接近中间值 $1/2$, 或者斜率变点的信号强度越大(即 $A_{\gamma\gamma}$ 越大), 则渐近分布的方差越小.

§4 数值模拟

在这一节, 将通过Monte Carlo模拟来评估估计量在有限样本情形下的表现. 对所有实验, 重复次数都为2000次. 数据由以下过程产生.

$$y_{i,t} = \mu_i^0 + \beta_i^0 t + \gamma_i^0 B_{T_1^0}(t) + u_{i,t}, \quad t = 1, \dots, T; \quad i = 1, \dots, N,$$

这里

$$u_{i,t} = h'_i F_t + e_{i,t}.$$

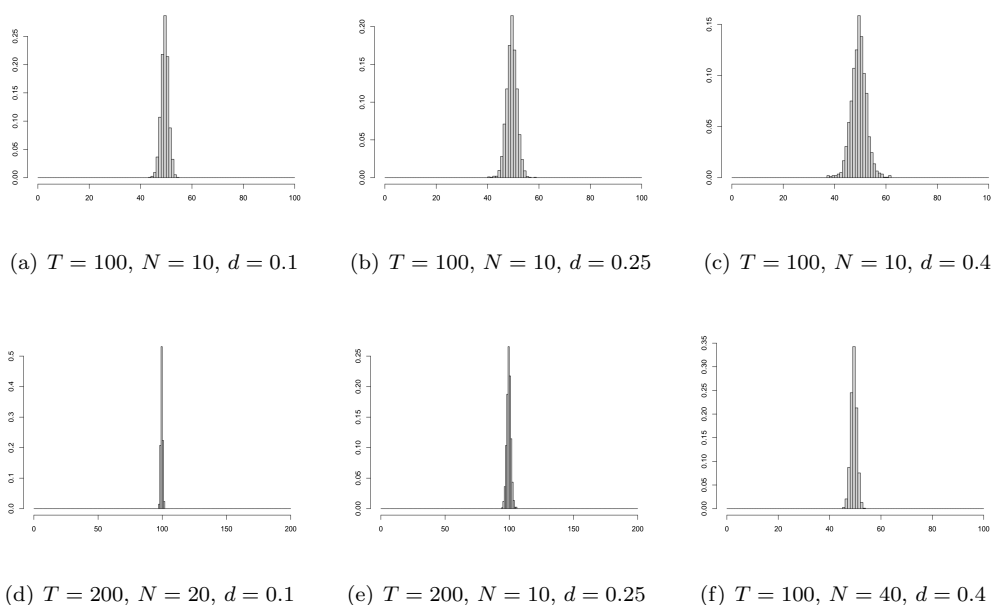


图 1 当所有的 $h_i = 0$, 且 $N = o(T^{3-2d})$ 时, \hat{T}_1 的直方图

设公因子 F_i 是一维的且满足 $F_t = 0.2F_{t-1} + w_t$, $F_0 = 0$, w_t 独立同分布于 $N(0, 1)$; 一维因子载荷 h_i 有两种情况: 若不存在横截面相依, 则对所有的 i 取 $h_i = 0$, 若公因子与变点的变化幅度存在交互效应, 即 $A_{H\gamma} \neq 0$, 设 h_i 由 $U(0, 1)$ 随机生成; 参考 McLeod 和 Hipel^[28] 以及 Hosking^[29] 的方法, 随机生成 N 个长相依时间序列 $\{e_{i,t}, t = 1, \dots, T\}$, $i = 1, \dots, N$, 使得 $(1-L)^d e_{i,t} = \varepsilon_{i,t}$, $\varepsilon_{i,t} = \sigma_i \eta_{i,t}$, 其中 σ_i 由 $U(1, 1.5)$ 随机生成, $\{\eta_{i,t}, t \geq 1, i \geq 1\}$ 独立同分布于 $N(0, 1)$, 记忆参数根据具体情况从 $d \in \{0.05, 0.1, 0.25, 0.4, 0.45\}$ 中选取. 为方便起见, 设定 $\mu_i^0 = \beta_i^0 = 0$, 因为这些参数的大小并不会影响 Monte Carlo 模拟的结果, 影响模拟结果的是斜率的变化幅度 γ_i^0 . 设 γ_i^0 由 $U(0, 0.3)$ 随机生成. 此外, 设定变点时刻的真值 $T_1^0 = 0.5T$, 以及 $\pi = 0.05$. 接下来, 用直方图的形式来评估 \hat{T}_1 在定理 3.1 中的表现.

图 1 展示了 $(T, N) \in \{(100, 10), (100, 40), (200, 10), (200, 20)\}$, $d \in \{0.1, 0.25, 0.4\}$, 所有的 $h_i = 0$ 时, \hat{T}_1 的直方图. 横向比较第一行可以看出估计量的估计误差随记忆参数 d 的增大而增加; 依次纵向比较可以看出固定 d 时估计误差随 T 和 N 的增大而减小, 这与定理 3.1(1) 的结论相符. 图 2 展示了 $(T, N) \in \{(200, 10), (200, 20), (400, 10), (400, 20)\}$, $d \in \{0.25, 0.4, 0.45\}$, 公因子与变点的变化幅度存在交互效应时, \hat{T}_1 的直方图, 这些 T, N, d 的设定是为了保证 $N = o(T^{2d})$ 或 $T^{2d} \asymp N$ 能够被满足. 同样通过横向和纵向的比较, 可以看出这与定理 3.1(2) 中的结论相符. 图 3 则展示了 $(T, N) \in \{(40, 100), (40, 200), (100, 100), (100, 200)\}$, $d \in \{0.05, 0.1, 0.25\}$, 公因子与变点的变化幅度存在交互效应时, \hat{T}_1 的直方图, 这些 T, N, d 的设定是为了保证 $T^{2d} = o(N)$ 能够被满足. 横向比较第一行可以看出变点估计量的估计误差与 d 无关, 纵向比较第一列可以看出估计误差与 N 亦无关, 纵向比较第二和三列则可看出变点估计量的估计误差与 T 有关, T 越大估计误差越小. 以上模拟结果与定理 3.1(3) 中的结论相符.

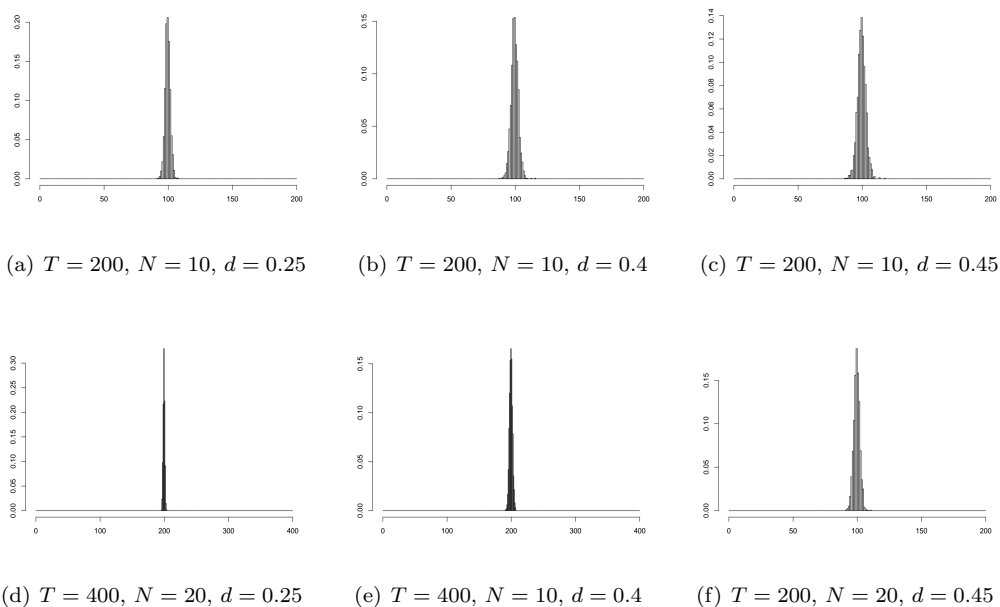


图 2 当 $A_{H\gamma} \neq 0$, 且 $N = o(T^{2d})$ 或 $T^{2d} \asymp N$ 时, \hat{T}_1 的直方图

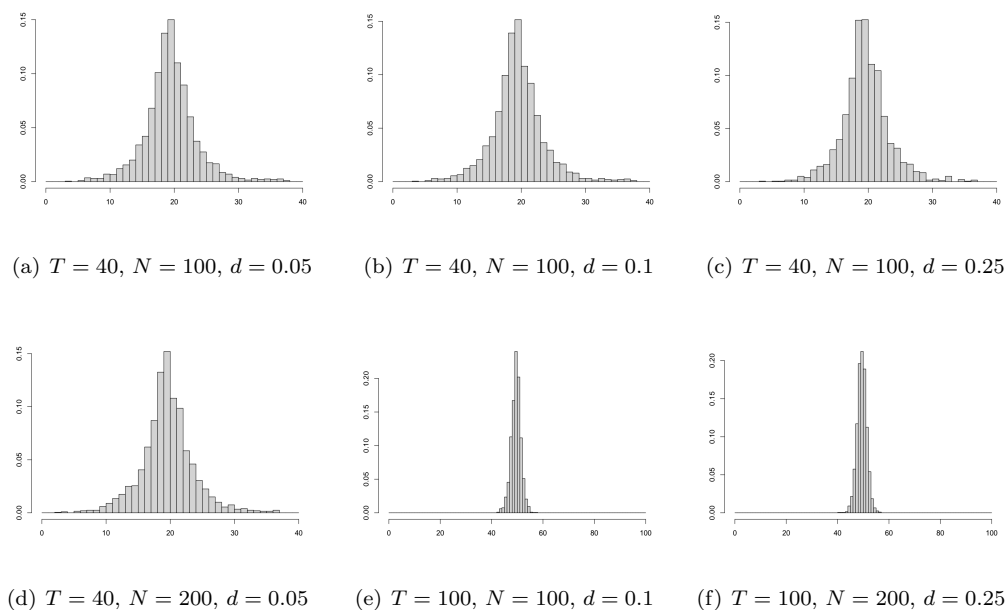


图 3 当 $A_{H\gamma} \neq 0$, 且 $T^{2d} = o(N)$ 时, \hat{T}_1 的直方图

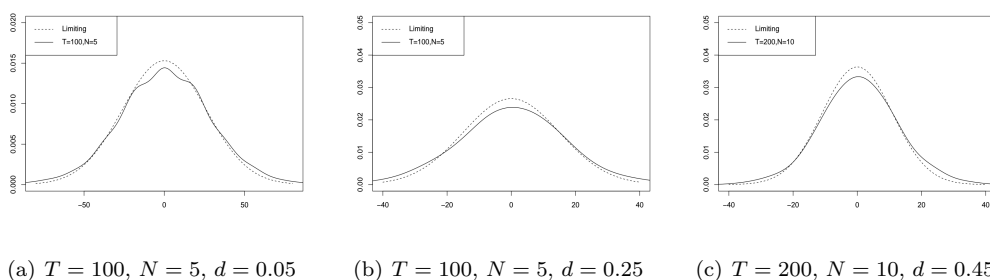


图 4 当所有的 $h_i = 0$, 且 $N = o(T^{3-2d})$ 时, $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布和极限分布

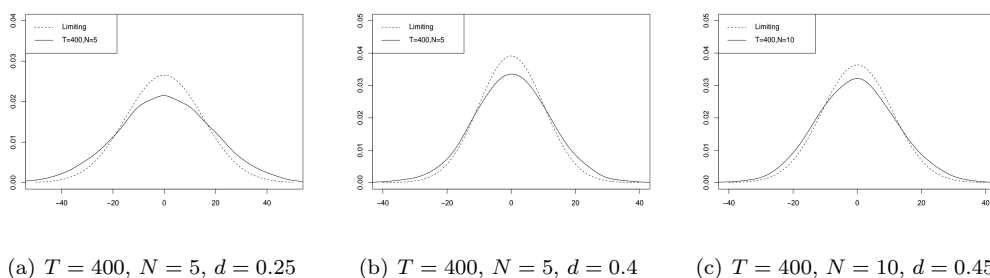
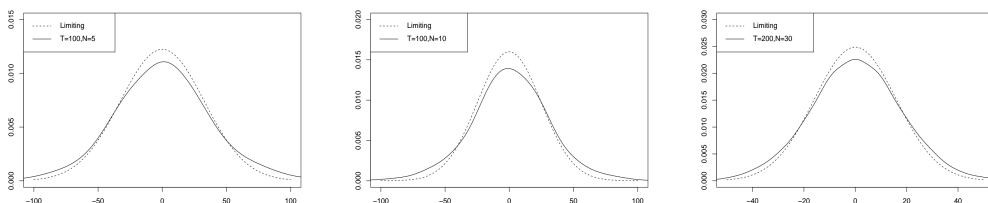


图 5 当 $A_{H\gamma} \neq 0$, 且 $N = o(T^{2d})$ 时, $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布和极限分布

最后比较估计量的有限样本分布与定理3.2的极限分布. 图4和图5展示 $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 在不存在横截面相依和公因子与变点的变化幅度存在交互效应这两种情况下, 估计量的有限样本分布与极限分布. 可以看出 d 越小, 有限样本分布与极限分布的逼近效果越好. 其次, 对于相同的 N 和 d , 不存在横截面相依的情况下有限样本分布更接近极限分布(即使 T 更小), 这说明公因子的存在会影响估计量的有限样本分布. 图6展示了当公因子与变点的变化幅度存在交互效应, 且 N/T^{2d} 收敛到不为零的常数时, $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布与极限分布. 可以看出, 两者的吻合程度较高. 最后, 图7展示了当公因子与变点的变化幅度存在交互效应且 $T^{2d} = o(N)$ 时, $T^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布与极限分布. 可以看出, 有限样本分布与极限分布的吻合度与 N 和 d 的关联较小, 但当 T 变大时, 有限样本分布更接近极限分布. 这些模拟结果与定理3.2(3)中的结论相符.

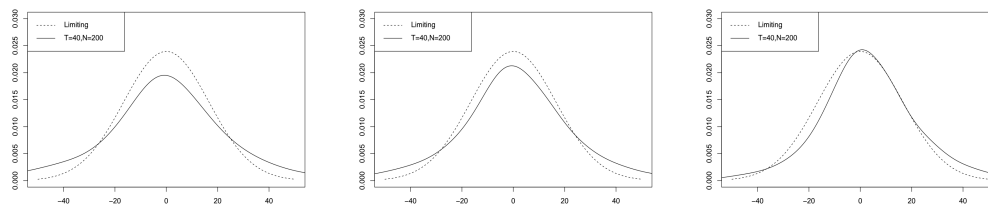
§5 结 论

本文研究了面板数据同时存在结构变点, 时间序列长相依以及横截面相依的变点估计问题. 用最小二乘的方法估计了变点发生的时刻, 并讨论了估计量的相合性, 收敛速度以及极限分布. 本文的结论表明在不存在横截面相依时, 变点时刻的估计量的收敛速度与记忆参数 d , 时间序列长度 T 以及个体数 N 有关, 且 d 越大收敛速度越慢; 当公因子与变点的变化幅度存在交互效应, 且 N 相对于 T 并没有足够大时, 收敛速度依然由 d , T 和 N 共同决定, 并随着 d 的增大而变慢; 然而当 $T^{2d} = o(N)$, 即 N 相对于 T 足够大时, 变点时刻估计量的收敛速度将仅与 T 有关. 最后通过 Monte Carlo 模拟比较了变点估计量的有限样本分布和极限分布, 并印证了本文的理论成果.

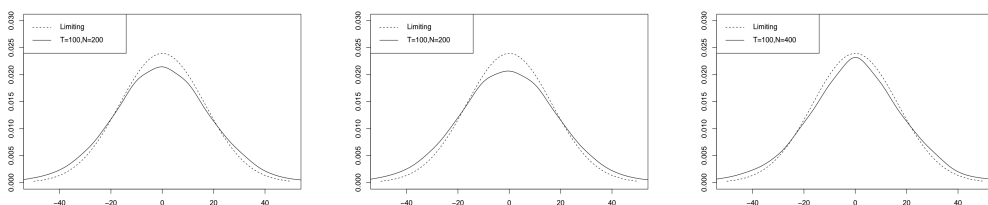


(a) $T = 100, N = 5, d = 0.1$ (b) $T = 100, N = 10, d = 0.25$ (c) $T = 200, N = 30, d = 0.4$

图 6 当 $A_{H\gamma} \neq 0$, 且 N/T^{2d} 收敛到不为零的常数时, $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布和极限分布



(a) $T = 40, N = 200, d = 0.05$ (b) $T = 40, N = 200, d = 0.25$ (c) $T = 40, N = 200, d = 0.45$



(d) $T = 100, N = 200, d = 0.05$ (e) $T = 100, N = 200, d = 0.25$ (f) $T = 100, N = 400, d = 0.45$

图 7 当 $A_{H\gamma} \neq 0$, 且 $T^{2d} = o(N)$ 时, $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0)$ 的有限样本分布和极限分布

§6 证明

本节给出定理的证明. 首先介绍关于 $\text{SSR}(\hat{T}_1) - \text{SSR}(T_1^0)$ 的一个分解式.

由于 $\mathbf{X}_{T_1^0} \mathbf{P}_{T_1^0} = \mathbf{X}_{T_1^0}$ 以及 $(\mathbf{I} - \mathbf{P}_{\hat{T}_1}) \mathbf{X}'_{\hat{T}_1} = 0$, 因此对于所有的 \hat{T}_1 都有

$$\begin{aligned} & \text{SSR}(\hat{T}_1) - \text{SSR}(T_1^0) \\ &= \text{tr}[\mathbf{Y}(\mathbf{I} - \mathbf{P}_{\hat{T}_1})\mathbf{Y}' - \mathbf{Y}(\mathbf{I} - \mathbf{P}_{T_1^0})\mathbf{Y}'] \\ &= \text{tr}[(\mathbf{\Pi} \mathbf{X}_{T_1^0} + \mathbf{U})(\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1})(\mathbf{X}'_{T_1^0} \mathbf{\Pi}' + \mathbf{U}')] \\ &= \text{tr}[\mathbf{\Pi} \mathbf{X}_{T_1^0} (\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1}) \mathbf{X}'_{T_1^0} \mathbf{\Pi}'] + 2\text{tr}[\mathbf{\Pi} \mathbf{X}_{T_1^0} (\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}'] + \text{tr}[\mathbf{U}(\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}'] \\ &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})(\mathbf{I} - \mathbf{P}_{\hat{T}_1})(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})' \mathbf{\Pi}'] + 2\text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})(\mathbf{I} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}'] \\ & \quad + \text{tr}[\mathbf{U}(\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}'] \leq 0. \end{aligned} \quad (2)$$

定义

$$\begin{aligned} S_{XX} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{\Pi}'], \\ S_{\hat{X}\hat{X}} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})(\mathbf{I} - \mathbf{P}_{\hat{T}_1})(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})' \mathbf{\Pi}'], \\ S_{XU} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1}) \mathbf{U}'], \\ S_{\hat{X}\hat{U}} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{\hat{T}_1})(\mathbf{I} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}'], \\ S_{UU} &= \text{tr}[\mathbf{U}(\mathbf{P}_{T_1^0} - \mathbf{P}_{T_1}) \mathbf{U}'], \\ S_{\hat{U}\hat{U}} &= \text{tr}[\mathbf{U}(\mathbf{P}_{T_1^0} - \mathbf{P}_{\hat{T}_1}) \mathbf{U}']. \end{aligned}$$

不等式(2)表明对于所有的 \hat{T}_1 , 均有

$$S_{\hat{X}\hat{X}} + 2S_{\hat{X}\hat{U}} + S_{\hat{U}\hat{U}} \leq 0. \quad (3)$$

定理3.1的证明 由式(1)中 \hat{T}_1 的定义, 可知

$$\hat{T}_1 = \arg \min_{T_1 \in [\pi T_1^0, (1-\pi)T_1^0]} [\text{SSR}(T_1) - \text{SSR}(T_1^0)],$$

这是因为 $\text{SSR}(T_1^0) = \text{tr}[\mathbf{Y}(\mathbf{I} - \mathbf{P}_{T_1^0})\mathbf{Y}']$ 是不依赖于 T_1 的. 此外根据式(2)

$$\text{SSR}(T_1) - \text{SSR}(T_1^0) = S_{XX} + 2S_{XU} + S_{UU}.$$

值得一提的是, 在后面的证明中, 仅给出 $T_1 \geq T_1^0$ 情形下的证明, 因为 $T_1 < T_1^0$ 情形下的证明是类似的. 接下来将分别研究 S_{XX} , S_{XU} 以及 S_{UU} , 并探究他们之间阶的大小关系.

首先来分析 S_{XX} . 容易看出, 在 $\lambda \in (0, 1)$ 下, 一致地有

$$\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1}) = \gamma(B_{T_1^0} - B_{T_1})' = \gamma(T_1 - T_1^0)\tilde{l}_b. \quad (4)$$

因此

$$\begin{aligned} S_{XX} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{\Pi}'] \\ &= |T_1 - T_1^0|^2 \text{tr}[\gamma \tilde{l}_b (\mathbf{I} - \mathbf{P}_{T_1}) \tilde{l}_b' \gamma'] = |T_1 - T_1^0|^2 (\gamma' \gamma) \tilde{l}_b (\mathbf{I} - \mathbf{P}_{T_1}) \tilde{l}_b' = |T_1 - T_1^0|^2 O(NT), \end{aligned} \quad (5)$$

这是因为 $\tilde{l}_b (\mathbf{I} - \mathbf{P}_{T_1}) \tilde{l}_b' = O(T)$ (见Perron和Zhu^[4], 第97页)以及 $\gamma' \gamma = O(N)$ (见假设3).

在假设2(1)下, 利用线性过程的泛函中心极限定理(见Phillips和Solo^[26]), 有

$$\begin{cases} T^{-1/2} \sum_{t=1}^T F_t \xrightarrow{d} \boldsymbol{\psi}(1)' \mathbf{W}_m(1), \\ T^{-3/2} \sum_{t=1}^T t F_t \xrightarrow{d} \boldsymbol{\psi}(1)' \int_0^1 r d\mathbf{W}_m(r), \\ T^{-3/2} \sum_{t=T_1+1}^T (t - T_1) F_t \Rightarrow \boldsymbol{\psi}(1)' \int_{\lambda}^1 (r - \lambda) d\mathbf{W}_m(r), \quad 0 < \lambda < 1, \end{cases}$$

这里 $\mathbf{W}_m(\cdot)$ 是一个 m 维的标准布朗运动. 在假设2(2)下, 利用长相依随机变量序列的泛函中心极

限定理(见Wang等作者^[25]中的定理2.1), 对于任意的 $i \geq 1$, 都有

$$\begin{cases} T^{-(1/2+d)} \sum_{t=1}^T e_{i,t} \xrightarrow{d} \kappa_i(d) B_d(1), \\ T^{-(3/2+d)} \sum_{t=1}^T t e_{i,t} \xrightarrow{d} \kappa_i(d) [B_d(1) - \int_0^1 B_d(r) dr] = \kappa_i(d) \int_0^1 r dB_d(r), \\ T^{-(3/2+d)} \sum_{t=T_1+1}^T (t - T_1) e_{i,t} \Rightarrow \kappa_i(d) \int_\lambda^1 (r - \lambda) dB_d(r), \quad 0 < \lambda < 1, \end{cases} \quad (6)$$

其中 $\kappa_i(d)^2 = \{\sigma_i^2 \Gamma(1 - 2d)\} / \{(1 + 2d)\Gamma(1 + d)\Gamma(1 - d)\}$. 此外, 易知在 $\lambda \in (0, 1)$ 上一致地有

$$\begin{cases} T^{-3} \sum_{t=T_1+1}^T (t - T_1)^2 \rightarrow \int_\lambda^1 (r - \lambda)^2 dr, \\ T^{-3} \sum_{t=T_1+1}^T (t - T_1)t \rightarrow \int_\lambda^1 (r - \lambda)r dr, \\ T^{-2} \sum_{t=T_1+1}^T (t - T_1) \rightarrow \int_\lambda^1 (r - \lambda) dr. \end{cases}$$

定义对角矩阵

$$D_T = \text{diag}(T, T^3, T^3).$$

基于上述结果和记号, 介绍下面的6个结论, 这些结论将帮助进一步分析 S_{XU} 和 S_{UU} .

(1) $D_T^{-1/2} X_{T_1} X'_{T_1} D_T^{-1/2}$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $O(1)$; 同样地, $D_T^{-1/2} X_{T_1^0} X'_{T_1^0} D_T^{-1/2}$ 的阶也是 $O(1)$; 见Perron和Zhu^[4], 第98页.

(2) $D_T^{-1/2} X_{T_1} F'$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $O_p(1)$; 同样地, $F X'_{T_1^0} D_T^{-1/2}$ 的阶也是 $O_p(1)$; 见Perron和Zhu^[4], 第105页.

(3) $D_T^{-1/2} X_{T_1} E_i$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $O_p(T^d)$; 同理, $E'_i X'_{T_1^0} D_T^{-1/2}$ 的阶也是 $O_p(T^d)$; 理由如下: 由式(6)可知

$$T^{-d} D_T^{-1/2} X_{T_1} E_i = \begin{pmatrix} T^{-(1/2+d)} \sum_{t=1}^T e_{i,t} \\ T^{-(3/2+d)} \sum_{t=1}^T t e_{i,t} \\ T^{-(3/2+d)} \sum_{t=T_1+1}^T (t - T_1) e_{i,t} \end{pmatrix} = O_p(1), \quad (7)$$

同理可得 $T^{-d} D_T^{-1/2} X_{T_1^0} E_i = O_p(1)$.

(4) $D_T^{-1/2} (X_{T_1^0} - X_{T_1}) F'$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $|T_1 - T_1^0| O_p(T^{-1})$; 见Perron和Zhu^[4], 第105页.

(5) $D_T^{-1/2} (X_{T_1^0} - X_{T_1}) E_i$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $|T_1 - T_1^0| O_p(T^{-1+d})$. 理由如下: 首先, 由于 $X_{T_1^0} - X_{T_1}$ 的前两行为零, 所以仅需考虑第三行. 此外

$$\begin{aligned} T^{-(1/2+d)} (B_{T_1^0} - B_{T_1})' E_i &= T^{-(1/2+d)} (T_1 - T_1^0) \tilde{e}_b E_i \\ &= T^{-(1/2+d)} \sum_{t=T_1^0+1}^{T_1} (t - T_1^0) e_{i,t} + T^{-(1/2+d)} (T_1 - T_1^0) \sum_{t=T_1+1}^T e_{i,t} \\ &= |T_1 - T_1^0| O_p(1). \end{aligned} \quad (8)$$

这说明 $D_T^{-1/2} (X_{T_1^0} - X_{T_1}) E_i$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $|T_1 - T_1^0| O_p(T^{-1+d})$.

(6) $D_T^{-1/2} (X_{T_1} X'_{T_1} - X_{T_1^0} X'_{T_1^0}) D_T^{-1/2}$ 的阶在 $\lambda \in (0, 1)$ 上一致地为 $|T_1 - T_1^0| O(T^{-1})$; 见Perron和Zhu^[4], 第98-99页.

接下来分析 S_{XU} . 首先由式(4)可知

$$\begin{aligned} S_{XU} &= \text{tr}[\Pi (X_{T_1^0} - X_{T_1}) (I - P_{T_1}) U'] \\ &= \text{tr}[\Pi (X_{T_1^0} - X_{T_1}) (I - P_{T_1}) F' H'] + \text{tr}[\Pi (X_{T_1^0} - X_{T_1}) (I - P_{T_1}) E'] \\ &= |T_1 - T_1^0| \tilde{e}_b (I - P_{T_1}) F' H' \gamma + |T_1 - T_1^0| \tilde{e}_b (I - P_{T_1}) E' \gamma. \end{aligned} \quad (9)$$

考虑式(9)中的第一项, 根据Kim^[17]中引理A.4的结论有

$$T^{-1/2} \tilde{e}_b (I - P_{T_1}) F' \xrightarrow{d} \zeta_F, \quad (10)$$

其中 $\zeta_F \sim N_m(0, \lambda^0(1 - \lambda^0)\psi(1)\psi(1)'/4)$. 因此结合假设4有

$$\tilde{t}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma = O_p(T^{1/2}N). \tag{11}$$

对于式(9)中的第二项, $\tilde{t}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{E}'\gamma = \tilde{t}_b\mathbf{E}'\gamma - \tilde{t}_b\mathbf{P}_{T_1}\mathbf{E}'\gamma$. 因为 \tilde{t}_b 中的元素都不会比 ι 中的元素大, 因此可以推导出 $\tilde{t}_b\mathbf{E}'\gamma$ 的阶不会高于(其实等于) $O(T^{1/2+d}N^{1/2})$, 这是因为

$$\text{Var}(\iota'\mathbf{E}'\gamma) = \text{Var}\left(\sum_{i=1}^N \iota' E_i \gamma_i\right) = \sum_{i=1}^N \gamma_i^2 \text{Var}(\iota' E_i) = O(T^{1+2d}N). \tag{12}$$

接下来, 很容易推出

$$\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{E}'\gamma = O_p(T^dN^{1/2}), \tag{13}$$

理由如下: $\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{E}'\gamma$ 是由 $T^{-1/2}\iota'\mathbf{E}'\gamma$, $T^{-3/2}\tau'\mathbf{E}'\gamma$ 和 $T^{-3/2}B'_{T_1}\mathbf{E}'\gamma$ 组成, 而 $T^{-1}\tau'$ 和 $T^{-1}B'_{T_1}$ 中的元素也都不会比 ι 中的元素大, 因此 $\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{E}'\gamma$ 的阶由 $T^{-1/2}\iota'\mathbf{E}'\gamma$ 决定, 而后者的阶为 $O_p(T^dN^{1/2})$ (根据式(12)). 由此可推得

$$\begin{aligned} \tilde{t}_b\mathbf{P}_{T_1}\mathbf{E}'\gamma &= \tilde{t}_b\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{E}'\gamma \\ &= O(T^{1/2})O(1)O_p(T^dN^{1/2}), \end{aligned}$$

其中 $\tilde{t}_b\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2} = O(T^{1/2})$ (见 Perron 和 Zhu^[4], 第104页). 于是

$$\tilde{t}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{E}'\gamma = O_p(T^{1/2+d}N^{1/2}). \tag{14}$$

由式(9)可知, 若对所有 i 都有 $h_i = 0$, 则 S_{XU} 仅剩其中的第二项; 若不然, 还需要比较式(11)和式(14)阶的大小, 即比较 T^{2d} 和 N 的相对大小. 若 $N = o(T^{2d})$, 则 S_{XU} 的阶将由式(9)中的第二项决定; 若 $T^{2d} = o(N)$, 则 S_{XU} 的阶将由式(9)中的第一项决定; 若 T^{2d} 和 N 同阶, 则 S_{XU} 的阶将由式(9)整体决定. 因此在 $\lambda \in (0, 1)$ 上一致地有

$$S_{XU} = \begin{cases} |T_1 - T_1^0|O_p(T^{1/2+d}N^{1/2}), & h_i = 0, \\ |T_1 - T_1^0|O_p(T^{1/2+d}N^{1/2}), & A_{H\gamma} \neq 0, N = o(T^{2d}) \text{ 或 } N \asymp T^{2d}, \\ |T_1 - T_1^0|O_p(T^{1/2}N), & A_{H\gamma} \neq 0, T^{2d} = o(N). \end{cases} \tag{15}$$

最后来分析 S_{UU} . 写

$$\begin{aligned} S_{UU} &= \text{tr}[\mathbf{U}(\mathbf{P}_{T_1^0} - \mathbf{P}_{T_1})\mathbf{U}'] \\ &= \text{tr}[\mathbf{U}[\mathbf{X}'_{T_1^0}(\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0})^{-1}\mathbf{X}_{T_1^0} - \mathbf{X}'_{T_1}(\mathbf{X}_{T_1}\mathbf{X}'_{T_1})^{-1}\mathbf{X}_{T_1}]\mathbf{U}'] \\ &= \text{tr}[\mathbf{U}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})'\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{U}'] \\ &\quad + \text{tr}[\mathbf{U}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}(\mathbf{X}_{T_1}\mathbf{X}'_{T_1} - \mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0})\mathbf{D}_T^{-1/2} \\ &\quad (\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{U}'] \\ &\quad + \text{tr}[\mathbf{U}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})\mathbf{U}'] \\ &= R_1 + R_2 + R_3. \end{aligned}$$

进一步地, 写

$$\begin{aligned} R_1 &= \text{tr}[\mathbf{U}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})'\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{U}'] \\ &= \text{tr}[(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{U}'\mathbf{U}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})'\mathbf{D}_T^{-1/2}] \\ &= \text{vec}((\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{X}'_{T_1^0}\mathbf{D}_T^{-1/2})^{-1})'\text{vec}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1^0}\mathbf{U}'\mathbf{U}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})'\mathbf{D}_T^{-1/2}), \end{aligned}$$

$$\begin{aligned}
 & \text{其中 } \text{vec}((\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{X}_{T_1^0}' \mathbf{D}_T^{-1/2})^{-1}) = O(1), \\
 & \text{vec}(\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{U}' \mathbf{U} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}) \\
 & = \text{vec}(\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{F}' \mathbf{H}' \mathbf{H} \mathbf{F} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}) \\
 & \quad + \text{vec}(\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{F}' \mathbf{H}' \mathbf{E} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}) \\
 & \quad + \text{vec}(\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{E}' \mathbf{H} \mathbf{F} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}) + \text{vec}(\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{E}' \mathbf{E} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}) \\
 & =: r_{11} + r_{12} + r_{13} + r_{14}.
 \end{aligned}$$

若 $A_{H\gamma} \neq 0$, 由 $\mathbf{H}'\mathbf{H} = O(N)$, $\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{F}' = O_p(1)$ 以及 $\mathbf{F}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2} = |T_1^0 - T_1| O_p(T^{-1})$,

与式(13)的推导过程类似, 易知 $\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{E}' \mathbf{H} = O_p(T^d N^{1/2})$. 结合式(8)和(12), 同理可得 $\mathbf{D}_T^{-1/2} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1}) \mathbf{E}' \mathbf{H} = |T_1^0 - T_1| O_p(T^{-1+d} N^{1/2})$. 因此

$$\begin{cases} r_{11} = |T_1^0 - T_1| O_p(T^{-1} N), \\ r_{12} = |T_1^0 - T_1| O_p(T^{-1+d} N^{1/2}), \\ r_{13} = |T_1^0 - T_1| O_p(T^{-1+d} N^{1/2}). \end{cases}$$

对于 r_{14} 来说, 注意到

$$\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} \mathbf{E}' \mathbf{E} (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2} = \sum_{i=1}^N \mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} E_i E_i' (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2},$$

其中 $\{\mathbf{D}_T^{-1/2} \mathbf{X}_{T_1^0} E_i E_i' (\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{D}_T^{-1/2}, i \geq 1\}$ 是相互独立的. 所以由式(7)和(8)可推得

$$r_{14} = |T_1^0 - T_1| O_p(T^{-1+2d} N).$$

综合以上结论可得

$$R_1 = |T_1^0 - T_1| O_p(T^{-1+2d} N).$$

同理可得

$$R_2 = |T_1^0 - T_1| O_p(T^{-1+2d} N), \quad R_3 = |T_1^0 - T_1| O_p(T^{-1+2d} N).$$

因此 $S_{UU} = |T_1 - T_1^0| O_p(T^{-1+2d} N)$. 值得注意的是, 若对所有的 i 都有 $h_i = 0$, 则 $r_{11} = r_{12} = r_{13} = 0$, 因此 $R_1 = r_{14} = |T_1^0 - T_1| O_p(T^{-1+2d} N)$. 对 R_2 和 R_3 亦有相同的结论. 因此 $S_{UU} = |T_1 - T_1^0| O_p(T^{-1+2d} N)$. 所以在 $\lambda \in (0, 1)$ 下一致地有

$$S_{UU} = |T_1 - T_1^0| O_p(T^{-1+2d} N). \tag{16}$$

综合(5), (15)和(16), 有如下结论.

$$\begin{cases} S_{XX} = |T_1 - T_1^0|^2 O(TN), \\ S_{XU} = \begin{cases} |T_1 - T_1^0| O_p(T^{1/2+d} N^{1/2}), & h_i = 0, \\ |T_1 - T_1^0| O_p(T^{1/2+d} N^{1/2}), & A_{H\gamma} \neq 0, N = o(T^{2d}) \text{ 或 } N \asymp T^{2d}, \\ |T_1 - T_1^0| O_p(T^{1/2} N), & A_{H\gamma} \neq 0, T^{2d} = o(N), \end{cases} \\ S_{UU} = |T_1 - T_1^0| O_p(T^{-1+2d} N). \end{cases}$$

用反证法论证. (1) 当对所有的 i 都有 $h_i = 0$ 时, 假设 $|T_1 - T_1^0|$ 的阶为 $k_T O_p(T^{-1/2+d} N^{-1/2})$, 其中 k_T 是关于 T 的正常数序列且是递增发散到无穷大的. 则有 $S_{\hat{X}\hat{X}} = O_p(k_T^2 T^{2d})$, $S_{\hat{X}\hat{U}} = O_p(k_T T^{2d})$ 以及 $S_{\hat{U}\hat{U}} = O_p(k_T T^{-3/2+3d} N^{1/2})$. 若 $N = o(T^{3-2d})$ 或 $N \asymp T^{3-2d}$, 则 $S_{\hat{X}\hat{X}}$ 的阶将会严格地大于其余的两项的阶, 这样式(3)将不能以概率1成立. 所以当对所有的 i 都有 $h_i =$

0时, $|\hat{T}_1 - T_1|$ 的阶最高不超过 $O_p(T^{-1/2+d}N^{-1/2})$. (2) 当 $A_{H\gamma} \neq 0$ 时, 需要分类讨论. 若 $N = o(T^{2d})$ 或 $N \asymp T^{2d}$, 假设 $|\hat{T}_1 - T_1|$ 的阶为 $k_T O_p(T^{-1/2+d}N^{-1/2})$, 和前面的论证过程类似, 会导致式(3)将不能以概率1成立. 因此 $|\hat{T}_1 - T_1|$ 的阶最高不超过 $O_p(T^{-1/2+d}N^{-1/2})$. 若 $T^{2d} = o(N)$, 假设 $|\hat{T}_1 - T_1|$ 的阶为 $k_T O_p(T^{-1/2})$, 同样会导致式(3)将不能以概率1成立. 因此 $|\hat{T}_1 - T_1|$ 的阶最高不超过 $O_p(T^{-1/2})$.

定理3.2的证明 先证明定理3.2(1). 给定正数 C , 定义集合

$$D(C) = \{T_1 : |T_1 - T_1^0| < CT^{-1/2+d}N^{-1/2}\},$$

以及

$$m_T = N^{1/2}T^{1/2-d}|T_1 - T_1^0|.$$

推导极限分布需分析 $\arg \min_{T_1 \in D(C)} [\text{SSR}(T_1) - \text{SSR}(T_1^0)]$. 对于任意的 $T_1 \in D(C)$, 都有 $|T_1 - T_1^0| = O(T^{-1/2+d}N^{-1/2})$. 因此

$$S_{XX} = |T_1 - T_1^0|^2 O(TN) = O(T^{2d}), \quad S_{XU} = |T_1 - T_1^0| O_p(T^{1/2+d}N^{1/2}) = O_p(T^{2d})$$

以及 $S_{UU} = |T_1 - T_1^0| O_p(T^{-1+2d}N) = O_p(T^{-3/2+3d}N^{1/2})$. 那么

$$\begin{aligned} \arg \min_{T_1 \in D(C)} [\text{SSR}(T_1) - \text{SSR}(T_1^0)] &= \arg \min_{T_1 \in D(C)} [S_{XX} + 2S_{XU} + S_{UU}]/T^{2d} \\ &= \arg \min_{T_1 \in D(C)} [S_{XX}/T^{2d} + 2S_{XU}/T^{2d} + o_p(1)]. \end{aligned}$$

这里 S_{UU}/T^{2d} 是渐近可忽略的, 这是因为定理假设中有 $N = o(T^{3-2d})$ 或 $N = o(T^{2d})$. 所以下面仅需考虑 S_{XX}/T^{2d} 和 $2S_{XU}/T^{2d}$.

首先, 对于 S_{XX}/T^{2d} 有

$$\begin{aligned} S_{XX}/T^{2d} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})' \mathbf{\Pi}'] / T^{2d} \\ &= |T_1 - T_1^0|^2 \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \tilde{\iota}_b' \gamma' \gamma / T^{2d} \\ &= m_T^2 (T^{-1} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \tilde{\iota}_b') (N^{-1} \gamma' \gamma) = m_T^2 ((1 - \lambda^0) \lambda^0 / 4) A_{\gamma\gamma} + o(1), \end{aligned} \tag{17}$$

最后这个等式成立是由于 $\lim_{T \rightarrow \infty} T^{-1} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \tilde{\iota}_b' = (1 - \lambda^0) \lambda^0 / 4$ (见Kim^[17], 第325页)以及假设 $\lim_{N \rightarrow \infty} N^{-1} \gamma' \gamma = A_{\gamma\gamma}$.

接下来分析 S_{XU}/T^{2d} . 由式(9)可知, 若对所有的 i 都有 $h_i = 0$, 或者当 $A_{H\gamma} \neq 0$ 且 $N = o(T^{2d})$, 则 S_{XU} 的阶均由(9)中的第二项决定. 所以

$$\begin{aligned} S_{XU}/T^{2d} &= \text{tr}[\mathbf{\Pi}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{U}'] / T^{2d} \\ &= [|T_1 - T_1^0| \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \mathbf{E}' \gamma (1 + o_p(1))] / T^{2d} \\ &= m_T (N^{-1/2} T^{-(1/2+d)} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \mathbf{E}' \gamma) (1 + o_p(1)). \end{aligned} \tag{18}$$

由假设2(2), $(1 - L)^d e_{i,t} = \varepsilon_{i,t}$, $\varepsilon_{i,t} = \sigma_i \eta_{i,t}$, $\eta_{i,t} \sim \text{iid}(0, 1)$. 记 $\eta_i = (\eta_{i,1}, \dots, \eta_{i,T})'$, 则有

$$\begin{aligned} N^{-1/2} T^{-(1/2+d)} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) \mathbf{E}' \gamma &= N^{-1/2} \sum_{i=1}^N T^{-(1/2+d)} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) E_i \gamma_i \\ &= N^{-1/2} \sum_{i=1}^N C_i Q_{i,T}, \end{aligned} \tag{19}$$

这里

$$\begin{cases} C_i = \sigma_i \gamma_i, \\ Q_{i,T} = T^{-(1/2+d)} \tilde{\iota}_b'(\mathbf{I} - \mathbf{P}_{T_1}) (1 - L)^{-d} \eta_i. \end{cases}$$

令 $\tilde{\eta}_i = E_i/\sigma_i = (1-L)^{-d}\eta_i$, 易知 $\tilde{\eta}_i$ 是一个长相依过程. 对于 $Q_{i,T}$, 由连续映射定理, 参照式(6), 在一系列复杂的计算之后, 可得 $Q_{i,T}$ 的极限分布 Q_i 的具体表达式

$$\begin{aligned} Q_{i,T} &= T^{-(1/2+d)}\tilde{t}_b(\mathbf{I} - \mathbf{P}_{T_1})\tilde{\eta}_i \\ &= T^{-(1/2+d)}\tilde{t}_b\tilde{\eta}_i - T^{-(1/2+d)}\tilde{t}_b\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2}(\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\mathbf{X}'_{T_1}\mathbf{D}_T^{-1/2})^{-1}\mathbf{D}_T^{-1/2}\mathbf{X}_{T_1}\tilde{\eta}_i \\ &\xrightarrow{d}\tilde{\kappa}(d)\left[\int_{\lambda^0}^1 dB_d(r) - \left(\frac{\lambda^0 - 1}{2}, \frac{3(1 - \lambda^0)}{2\lambda^0}, \frac{3(2\lambda^0 - 1)}{2\lambda^0(1 - \lambda^0)}\right) \begin{pmatrix} \int_0^1 dB_d(r) \\ \int_0^1 r dB_d(r) \\ \int_{\lambda^0}^1 (r - \lambda^0) dB_d(r) \end{pmatrix}\right] \\ &= \tilde{\kappa}(d)\left[\int_0^{\lambda^0} \frac{\lambda^0 - (\lambda^0)^2 - 3r + 3r\lambda^0}{2\lambda^0} dB_d(r) + \int_{\lambda^0}^1 \frac{\lambda^0(2 + \lambda^0 - 3r)}{2(1 - \lambda^0)} dB_d(r)\right] \\ &= \tilde{\kappa}(d)\zeta_{\tilde{\eta}} =: Q_i, \end{aligned}$$

这里 $\tilde{\kappa}(d)^2 = \{\Gamma(1 - 2d)\}/\{(1 + 2d)\Gamma(1 + d)\Gamma(1 - d)\}$. 容易看出 $E(Q_i) = 0$, $\text{Var}(Q_i) = \tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2$. 由假设2(3)知 $Q_{i,T}^2$ 关于 T 一致可积. 则 $\text{Var}(Q_{i,T}) = E(Q_{i,T}^2) \rightarrow E(Q_i^2) = \tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2$. 此外假设3表明

$$\max_i \frac{C_i^2}{\sum_i C_i^2} = \max_i \frac{\sigma_i^2\gamma_i^2}{\gamma'\boldsymbol{\Sigma}_\varepsilon\gamma} \leq N^{-1} \frac{M \max_i \gamma_i^2}{N^{-1}\gamma'\boldsymbol{\Sigma}_\varepsilon\gamma} = O(N^{-1}).$$

因此根据联合中心极限定理(见Phillips和Moon^[27]中的定理3)有

$$N^{-1/2} \sum_{i=1}^N C_i Q_{i,T} \xrightarrow{d} N(0, \tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2 S_{\gamma\gamma}). \tag{20}$$

最后根据式(18)和(19)可知 $S_{XU}/(T^{2d}m_T) \xrightarrow{d} N(0, \tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2 S_{\gamma\gamma})$.

综上所述, 由 $\arg \max / \arg \min$ 泛函的连续映射定理(见Kim和Pollard^[30])及一些简单的计算, 可知

$$\begin{aligned} T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0) &= m_T^* = \arg \min_{m_T \in D(C)} [S_{XX}/T^{2d} + 2S_{XU}/T^{2d} + o_p(1)] \\ &\xrightarrow{d} N\left(0, \frac{16\tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2 S_{\gamma\gamma}}{(1 - \lambda^0)^2(\lambda^0)^2 A_{\gamma\gamma}^2}\right). \end{aligned}$$

由于 $T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0) = O_p(1)$, 所以存在足够大的 C 使得 $\hat{T}_1 \in D(C)$ 的概率任意地接近1, 从而定理3.2(1)得证.

接下来证明定理3.2(2). 给定正数 C , 定义集合

$$D(C) = \{T_1 : |T_1 - T_1^0| < CT^{-1/2+d}N^{-1/2}\},$$

以及

$$m_T = N^{1/2}T^{1/2-d}|T_1 - T_1^0|.$$

$\forall T_1 \in D(C)$, 都有 $|T_1 - T_1^0| = O(T^{-1/2+d}N^{-1/2})$. 因此 $S_{XX} = O(T^{2d})$, $S_{XU} = O(T^{2d})$ 以及 $S_{UU} = O_p(T^{-3/2+3d}N^{1/2})$. 那么

$$\arg \min_{T_1 \in D(C)} [\text{SSR}(T_1) - \text{SSR}(T_1^0)] = \arg \min_{T_1 \in D(C)} [S_{XX}/T^{2d} + 2S_{XU}/T^{2d} + o_p(1)],$$

S_{XU}/T^{2d} 是渐近可忽略的是因为 $N/T^{2d} \rightarrow c \neq 0$.

对于 S_{XX}/T^{2d} , 根据式(17)有 $S_{XX}/T^{2d} = m_T^2((1 - \lambda^0)\lambda^0/4)A_{\gamma\gamma} + o(1)$.

分析 S_{XU}/T^{2d} . 由式(9)可知, 当 $A_{H\gamma} \neq 0$ 且 $N/T^{2d} \rightarrow c \neq 0$ 时, S_{XU} 的阶由式(9)整体决定.

那么

$$\begin{aligned} S_{XU}/T^{2d} &= \text{tr}[\mathbf{II}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{U}']/T^{2d} \\ &= [|T_1 - T_1^0|\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma + |T_1 - T_1^0|\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{E}'\gamma]/T^{2d} \\ &= m_T(N^{-1/2}T^{-(1/2+d)}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma + N^{-1/2}T^{-(1/2+d)}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{E}'\gamma) \\ &= m_T(\sqrt{N/T^{2d}}N^{-1}T^{-1/2}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma + N^{-1/2}T^{-(1/2+d)}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{E}'\gamma). \end{aligned}$$

根据假设2(4)和假设4, 及式(10)和(20)可知

$$S_{XU}/(T^{2d}m_T) \xrightarrow{d} N(0, c\lambda^0(1 - \lambda^0)A'_{H\gamma}\psi(1)\psi(1)'A_{H\gamma}/4 + \tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2S_{\gamma\gamma}).$$

综上所述, 类似于定理3.2(1)中的证明, 有

$$\begin{aligned} T^{1/2-d}N^{1/2}(\hat{T}_1 - T_1^0) &= m_T^* = \arg \min_{m_T \in D(C)} [S_{XX}/T^{2d} + 2S_{XU}/T^{2d} + o_p(1)] \\ &\xrightarrow{d} N\left(0, \frac{4cA'_{H\gamma}\psi(1)\psi(1)'A_{H\gamma}}{\lambda^0(1 - \lambda^0)A_{\gamma\gamma}^2} + \frac{16\tilde{\kappa}(d)^2\sigma_{\tilde{\eta}}^2S_{\gamma\gamma}}{(1 - \lambda^0)^2(\lambda^0)^2A_{\gamma\gamma}^2}\right). \end{aligned}$$

定理3.2(2)得证.

最后证明定理3.2(3). 给定正数 C , 定义集合

$$D(C) = \{T_1 : |T_1 - T_1^0| < CT^{-1/2}\},$$

以及

$$m_T = T^{1/2}|T_1 - T_1^0|.$$

对于任意的 $T_1 \in D(C)$, 有 $|T_1 - T_1^0| = O(T^{-1/2})$. 因此, $S_{XX} = |T_1 - T_1^0|^2O(TN) = O(N)$, $S_{XU} = |T_1 - T_1^0|O_p(T^{1/2}N) = O_p(N)$ 以及 $S_{UU} = |T_1 - T_1^0|O_p(T^{-1+2d}N) = O_p(T^{-3/2+2d}N)$. 那么

$$\begin{aligned} \arg \min_{T_1 \in D(C)} [\text{SSR}(T_1) - \text{SSR}(T_1^0)] &= \arg \min_{T_1 \in D(C)} [S_{XX} + 2S_{XU} + S_{UU}]/N \\ &= \arg \min_{T_1 \in D(C)} [S_{XX}/N + 2S_{XU}/N + o_p(1)]. \end{aligned}$$

首先分析 S_{XX}/N . 容易看出

$$S_{XX}/N = m_T^2(T^{-1}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\tilde{u}_b'(N^{-1}\gamma'\gamma)) = m_T^2((1 - \lambda^0)\lambda^0/4)A_{\gamma\gamma} + o(1).$$

接下来分析 S_{XU}/N . 由式(9)可知, 当 $A_{H\gamma} \neq 0$ 且 $T^{2d} = o(N)$ 时, S_{XU} 的阶由式(9)中的第一项决定. 所以

$$\begin{aligned} S_{XU}/N &= \text{tr}[\mathbf{II}(\mathbf{X}_{T_1^0} - \mathbf{X}_{T_1})(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{U}']/N \\ &= [|T_1 - T_1^0|\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma(1 + o_p(1))]/N \\ &= m_T T^{-1/2}\tilde{u}_b(\mathbf{I} - \mathbf{P}_{T_1})\mathbf{F}'\mathbf{H}'\gamma(1 + o_p(1))/N. \end{aligned}$$

根据式(10)可知 $S_{XU}/(Nm_T) \xrightarrow{d} \zeta_F A_{H\gamma}$, 其中 $\zeta_F \sim N_m(0, \lambda^0(1 - \lambda^0)\psi(1)\psi(1)'/4)$.

综上所述, 类似于前面的证明可知

$$\begin{aligned} T^{1/2}(\hat{T}_1 - T_1^0) &= m_T^* = \arg \min_{m_T \in D(C)} [S_{XX}/N + 2S_{XU}/N + o_p(1)] \\ &\xrightarrow{d} N\left(0, \frac{4A'_{H\gamma}\psi(1)\psi(1)'A_{H\gamma}}{\lambda^0(1 - \lambda^0)A_{\gamma\gamma}^2}\right). \end{aligned}$$

定理3.2(3)得证.

参考文献:

- [1] Bai Jushan, Perron P. Estimating and testing linear models with multiple structural changes[J]. *Econometrica*, 1998, 66(1): 47-78.
- [2] Hansen B E. The new econometrics of structural change: Dating breaks in U.S. labor productivity[J]. *Journal of Economic Perspectives*, 2001, 15(4): 117-128.
- [3] Lee S, Seo M H, Shin Y. The lasso for high dimensional regression with a possible change point[J]. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2016, 78(1): 193-210.
- [4] Perron P, Zhu Xiaokang. Structural breaks with deterministic and stochastic trends[J]. *Journal of Econometrics*, 2010, 28(2): 275-290.
- [5] Horváth L, Kokoszka P. The effect of long-range dependence on change-point estimators[J]. *Journal of Statistical Planning and Inference*, 1997, 64: 57-81.
- [6] Jarušková D. Some problems with application of change-point detection methods to environmental data[J]. *Environmetrics*, 1997, 8469-483.
- [7] Kuan C, Hsu C. Change-point estimation of fractionally integrated processes[J]. *Journal of Time Series Analysis*, 1998, 19(6): 693-708.
- [8] Chang S Y, Perron P. Inference on a structural break in trend with fractionally integrated errors[J]. *Journal of Time Series Analysis*, 2016, 37(4): 555-574.
- [9] Perron P, Qu Zhongjun. Long memory and level shifts in the volatility of stock market return indices[J]. *Journal of Business & Economic Statistics*, 2009, 27(3): 369-396.
- [10] Shao Xiaofeng. A simple test of changes in mean in the possible presence of long-range dependence[J]. *Journal of Times Series Analysis*, 2011, 32(6): 598-606.
- [11] Iacone F, Leybourne S J, Taylor A M R. Testing the order of fractional integration of a time series in the possible presence of a trend break at an unknown point[J]. *Econometric Theory*, 2019, 35: 1201-1233.
- [12] Joseph L, Wolfson D. Estimation in the multi-path change-point problems[J]. *Communications in Statistics-Theory and Methods*, 1992, 21(4): 897-913.
- [13] Joseph L, Wolfson D. Maximum likelihood estimation in the multi-path change-point problem[J]. *Annals of the Institute of Statistical Mathematics*, 1993, 43(3): 511-530.
- [14] Joseph L, Vandal A, Wolfson D. Estimation in the multi-path change-point problem for correlated data[J]. *The Canadian Journal of Statistics*, 1996, 24(1): 37-54.
- [15] Bai Jushan. Common breaks in means and variance for panel data[J], *Journal of Econometrics*, 2010, 157(1): 78-92.
- [16] Horváth L, Hušková M. Change-point detection in panel data[J]. *Journal of Time Series Analysis*, 2012, 33(4): 631-648.
- [17] Kim D. Estimating a common deterministic time trend break in large panels with cross sectional dependence[J]. *Journal of Econometrics*, 2011, 164(2): 310-330.
- [18] Kim D. Common breaks in time trends for large panel data with a factor structure[J]. *The Econometrics Journal*, 2014, 17(3): 301-337.
- [19] Li Degui, Qian Junhui, Su Liangjun. Panel data models with interactive fixed effects and multiple structural breaks[J]. *Journal of the American Statistical Association*, 2016, 111(516): 1804-1819.
- [20] Qian Junhui, Su Liangjun. Shrinkage estimation of common breaks in panel data model via adaptive group fused Lasso[J]. *Journal of Econometrics*, 2016, 191: 86-109.

- [21] Baltagi B, Feng Qu, Kao C. Estimation of heterogeneous panels with structural breaks[J]. *Journal of Econometrics*, 2016, 191(1): 176-195.
- [22] Baltagi B, Feng Qu, Kao C. Structural changes in heterogeneous panels with endogenous regressors[J]. *Journal of Applied Econometrics*, 2019, 34(6): 883-892.
- [23] Westerlund J. Common breaks in means for cross-correlated fixed-T panel data[J]. *Journal of Time Series Analysis*, 2019, 40: 248-255.
- [24] Billingsley P. *Convergence of Probability Measures* (2nd edition)[M]. New York: John Wiley and Sons, 1999.
- [25] Wang Qiyang, Lin Yan-Xia, Gulati C M. Asymptotics for general fractionally integrated processes with applications to unit root tests[J]. *Econometric Theory*, 2003, 19(1): 143-164.
- [26] Phillips P C B, Solo V. Asymptotics for linear processes[J]. *The Annals of Statistics*, 1992, 20(2): 971-1001.
- [27] Phillips P C B, Moon H R. Linear regression limit theory for nonstationary panel data[J]. *Econometrica*, 1999, 67: 1057-1111.
- [28] McLeod A I, Hipel K W. Preservation of the rescaled adjusted range. 1: A reassessment of the Hurst phenomenon[J]. *Water Resources Research*, 1978, 14: 491-508.
- [29] Hosking J R M. Fractional differencing[J]. *Biometrika*, 1981, 68(1): 165-176.
- [30] Kim J, Pollard D. Cube root asymptotics[J]. *The Annals of Statistics*, 1990, 18(1): 191-219.

Change-point analysis for the slope in long-range dependent panel data

ZHU Xu, PANG Tian-xiao

(School of Mathematical Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract: Financial panel data often have the phenomena of structural change, long-range dependence and cross-sectional dependence simultaneously. Assuming that the length of time series is T , the number of individuals is N , the slope change occurs at the same time of all individuals, and each time series is long-range dependent (memory parameter $d \in (0, 0.5)$), and there is cross-sectional dependence between individuals due to the common factor structure. For this panel data model, we apply the least square method to estimate the date of slope change and the fraction of the slope change point, and study the asymptotic properties of the estimator when (T, N) jointly tends to infinity, including the consistency, convergence rate as well as the limiting distribution. Some interesting conclusions are drawn. In most cases, the convergence rate of the estimator of the slope change point slows down with the increase of the memory parameter d . However, when the time length T and the number of individuals N of the panel data satisfy the condition $T^{2d} = o(N)$, and there is an interaction effect between the common factor and the break magnitude of the change point, the convergence rate of the change point estimator is only related to T . Monte Carlo simulations are conducted to evaluate the finite-sample performance of the estimator, and our theoretical results are supported by the Monte Carlo simulations.

Keywords: panel data; long-range dependence; slope change point; least squares estimator

MR Subject Classification: 60F05; 62F12